

WO 01/97794 A2

STRUCTURE BASED DISCOVERY OF INHIBITORS OF MATRIPTASE
FOR THE TREATMENT OF CANCER AND OTHER CONDITIONS

CROSS REFERENCE TO RELATED APPLICATION

- 5 This application claims priority from U.S. Provisional Serial No. 60/213,073, filed June 21, 2000, and is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

1. Field of the Invention

- 10 The present invention relates to methods for diagnosing and treating conditions involving matriptase activity, particularly cancer. In particular, the invention is based on the elucidation of matriptase activity and involvement in the development of precancerous and cancerous conditions. More particularly, the invention relates to the design of bioassay testing methods for cancer diagnosis based
15 on the detection of matriptase activity, identification of compounds capable of inhibiting matriptase activity and design of cancer therapy protocols based on the inhibition of matriptase employing small molecule inhibitors.

2. Discussion of the Related Art

- 20 *Serine Proteases and Other Cancer Related Proteases.* Elevated proteolytic activity has been implicated in neoplastic progression. While the exact role(s) of proteolytic enzymes in the progression of tumor remains unclear, it seems that proteases may be involved in almost every step of the development and spread of cancer. A widely proposed view is that proteases contribute to the degradation of
25 extracellular matrix (ECM) and to tissue remodeling, and are necessary for cancer progression and metastasis. A wide array of ECM-degrading proteases has been discovered, the expression of some of which correlates with tumor progression. These include matrix metalloproteases (MMPS) family, plasmin/urokinase type plasminogen activator system and lysosomal proteases cathepsins D and B reviewed
30 by Mignatti *et al.*, *Physiol. Rev.* 73: 161-95 (1993). The plasmin/urokinase type plasminogen activator system is composed of plasmin, the major ECM-degrading protease; the plasminogen activator, uPA; the plasmin inhibitor $\alpha 2$ -anti-plasmin, the plasminogen activator inhibitors PAI-1 and PAI-2; and the cell membrane receptor for

uPA (uPAR) (Andreasen *et al.*, *Int. J. Cancer* 72: 1-22 (1997)).

The MMPs are a family of zinc-dependent enzymes with characteristic structures and catalytic properties. The plasmin/urokinase type plasminogen activator system and the 72-kDa gelatinase (MMP-2)/membrane-type MMP system have received the most attention for their potential roles in the process of progression of breast cancer and other carcinomas. However, both systems appear to require indirect mechanisms to recruit and activate the major ECM-degrading proteases on the surface of cancer cells. For example, uPA is produced *in vivo* (Nielson *et al.*, *Lab. Invest.* 74: 168-77 (1996); Pyke *et al.*, *Cancer Res.* 53: 1911-15 (1993); Polette *et al.*, *Virchows Arch.* 424: 641-45 (1994); and Okada *et al.*, *Proc. Natl. Acad. Sci. USA* 92: 2730-34 (1995)) in human breast carcinomas by myofibroblasts adjacent to cancer cells and must diffuse to the cancer cells for receptor-mediated activation and presentation on the surfaces of cancer cells. However, the uPA receptor (uPAR) is detected in macrophages that infiltrate tumor foci in ductal breast cancer. Somewhat analogously, the majority of the MMP family members, such as 72-kDa/Gelatinase A (MMP-2) (Lin *et al.*, *J. Biol. Chem.* 272: 9147-52 (1997)), stromelysin-3 (MMP-11) (Matsudaira, *J. Biol. Chem.* 262: 10035-38 (1987)), MTMMP (MMP-14), are expressed by fibroblastic cells of tumor stroma, or surrounding noncancerous tissues, or both. Indirect mechanisms of activation and recruitment of Gelatinase A in the close vicinity of the surfaces of cancer cells have been proposed, such that an unidentified cancer cell-derived membrane receptor(s) of Gelatinase A could serve as membrane anchor for Gelatinase A; cleaved MT-MMP from stroma cells could then diffuse to the surfaces of cancer cells to activate Gelatinase A. Matrilysin (MMP-7; Pump-1) appears to be the only MMP which is found predominantly in the epithelial cells.

The stromal origins of these well-characterized extracellular matrix-degrading proteases may suggest that cancer progression is an event which either depends entirely upon stromal-epithelial cooperation or which is controlled by some other unknown epithelial-derived proteases. Search for these epithelial-derived proteolytic systems that may interact with plasmin/urokinase type plasminogen activator system and/or with MMP family could provide a missing link in the understanding of malignant progression.

Matriptase was initially identified from T-47D human breast cancer cells as a major gelatinase with a migration rate between those of Gelatinase A (72-kDa, MMP-2) and Gelatinase B (92-kDa, MMP-9). It has been proposed to play a role in the metastatic invasiveness of breast cancer. (See U.S. Patent 5,482,848, the contents of which are incorporated herein by reference in their entirety.) The primary cleavage specificity of matriptase was identified to be arginine and lysine residues, similar to the majority of serine proteases, including trypsin and plasmin. In addition, matriptase, like trypsin, exhibits broad spectrum cleavage activity, and such activity is likely to contribute to its gelatinolytic activity. The trypsin-like activity of matriptase distinguishes it from Gelatinases A and B, which may cleave gelatin at glycine residues, the most abundant (almost one third) of amino acid residues in gelatin.

Kunitz-type serine protease inhibitors. Hepatocyte growth factor (HGF) activator inhibitor-1 (HAI-1) is a Kunitz-type serine protease inhibitor which is able to inhibit HGF activator, a blood coagulation factor XII-like serine protease. The mature form of this protease inhibitor has 478 amino acid residues, with a calculated molecular mass of 53,319. A putative transmembrane domain is located at its carboxyl terminus. HAI-1 contains two Kunitz domains (domain I spans residues 246-306; domain II spans residues 371 to 431) separated by a LDL receptor domain (residues 315 to 360). The presumed P1 residue of active-site cleft is likely to be arginine-260 in Kunitz domain I and lysine 385 in domain II by alignment with bovine pancreatic trypsin inhibitor (BPTI, aprotinin) and with other Kunitz-type inhibitors. Thus, HAI-1 has specificity against trypsin-type proteases. Although HGF activator is exclusively expressed by liver cells, HAI-1 was originally purified from the conditioned media of carcinoma cells as a 40-kDa fragment doublet, rather than the proposed, mature, membrane-bound, 53-kDa form (Shimomura *et al.*, *J. Biol. Chem.* 272: 6370-76 (1997)).

The protein inhibitors of serine proteases can be classified into at least 10 families, according to various schemes. Among them, serpins, such as maspin (Sheng *et al.*, *Proc. Natl. Acad. Sci. USA* 93: 11669-74 (1996)) and Kunitz-type inhibitors, such as urinary trypsin inhibitor (Kobayashi *et al.*, *Cancer Res.* 54: 844-49 (1994)) have been previously implicated in suppression of cancer progression. The Kunitz-type inhibitors form very tight, but reversible complexes with their target serine

proteases. The reactive sites of these inhibitors are rigid and can simulate optimal protease substrates. The interaction between a serine protease and a Kunitz-type inhibitor depends on complementary, large surface areas of contact between the protease and inhibitor. The inhibitory activity of the recovered Kunitz-type inhibitor
5 from protease complexes can always be reconstituted. The Kunitz-type inhibitors may be cleaved by cognate proteases, but such cleavage is not essential for their inhibitory activity. In contrast, serpin-type inhibitors also form tight, stable complexes with proteases; in most of cases these complexes are even more stable than those containing Kunitz-type inhibitors. Cleavage of serpins by proteases is necessary
10 for their inhibition, and serpins are always recovered in a cleaved, inactive form from protease reactions.

Thus, serpins are considered to be suicide substrate inhibitors, and their inhibitory activity will be lost after encounters with proteases. The suicide nature of serpin inhibitors may result in regulation of proteolytic activity *in vivo* by direct
15 removal of unwanted proteases via other membrane-bound endocytic receptors (in the case of uPA inhibitors). However, the Kunitz type inhibitors may simply compete with physiological substrates (such as ECM components), and in turn, reduce their availability for proteolysis. These differences may result in different mechanisms whereby these proteases perform their roles in ECM-degradation and cancer
20 progression.

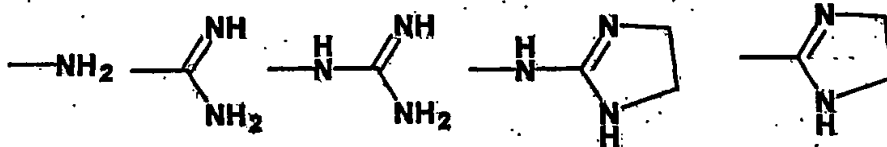
It has previously been disclosed that a soybean-derived compound known as Bowman-Birk inhibitor (BBI, from Sigma) may have anti-cancer activity by preventing tumor initiation and progression in model systems. However, due to its size and non-synthetic nature, use of BBI in therapeutic methods based on inhibition
25 of matriptase activity can be cumbersome and expensive.

For their stability and permeability, small molecule inhibitors provide valuable tools on functional studies in various systems. It is highly desirable to provide small molecules that can inhibit matriptase activity. Based on the potential effects thereof on carcinoma progression, cell migration, proliferation and apoptosis, it would be
30 beneficial if compounds could be obtained which selectively block the proteolytic activity of matriptase. More specifically, it would be beneficial if compounds could be obtained which antagonize cascade protease activators produced by cleavage of

matriptase substrates. Such compounds have significant therapeutic potential, in particular for treatment of cancer and other conditions involving carcinoma progression and migration and abnormal cell differentiation and proliferation. Compounds having improved selectivity, solubility and stability are particularly desirable.

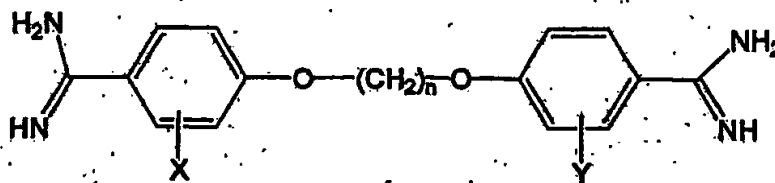
3. Summary and Objects of the Invention

In one embodiment, the present invention provides a method of inhibiting carcinoma progression wherein matriptase plays a role in a subject in need of such inhibition comprising administering to a subject an effective amount of a compound comprising two groups capable of being positively charged at physiological pH, which are the same or different. The groups are linked by a chemical group having a length of between 5 and 30 Å, and preferably between 15 and 24 Å. The groups are preferably selected from the following groups:

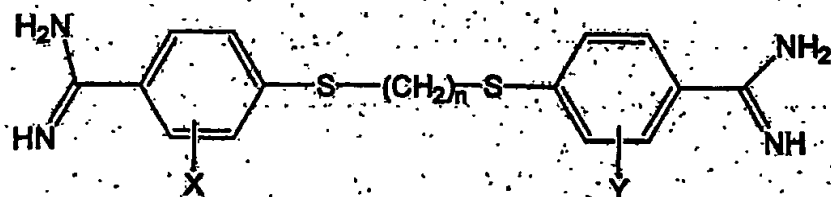


Preferred compounds include the compounds of Examples I through VI which have the structures below, wherein X and Y can be any substituents.

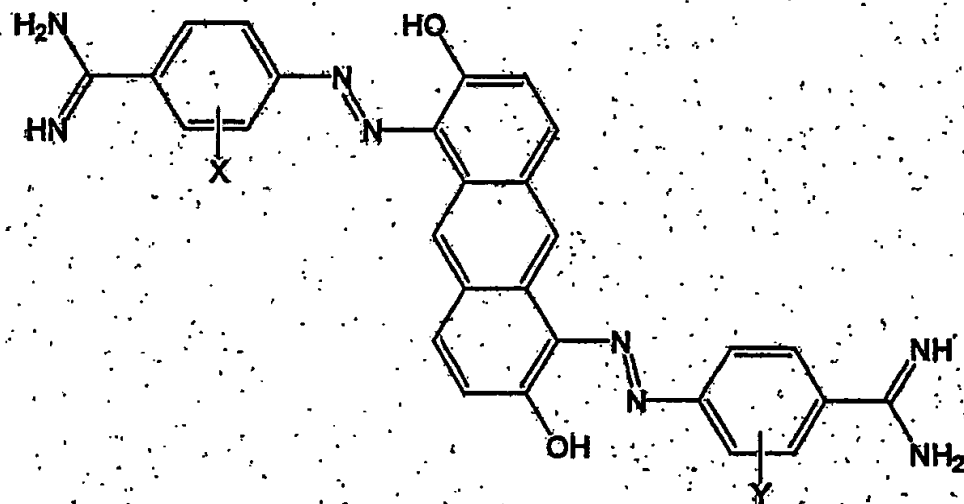
Example I

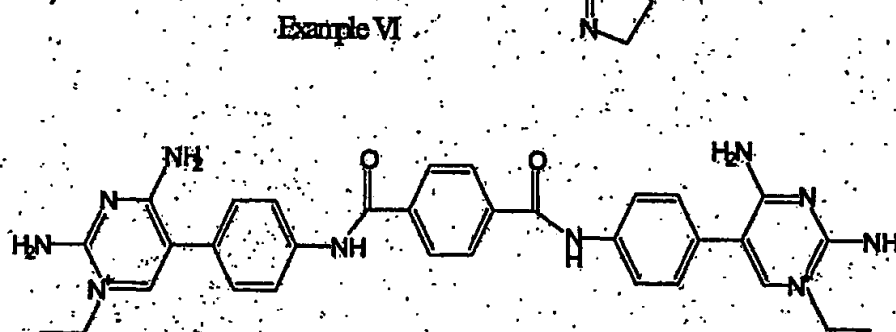
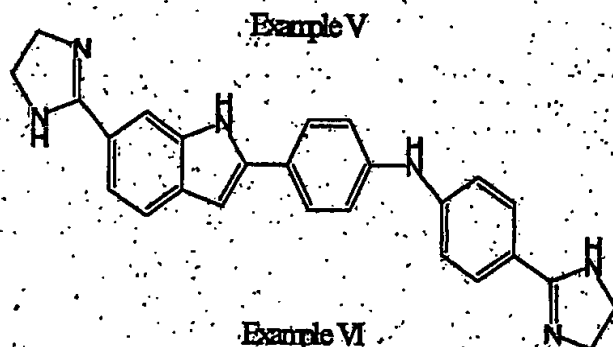
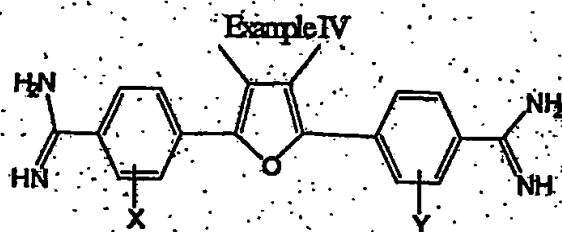


Example II

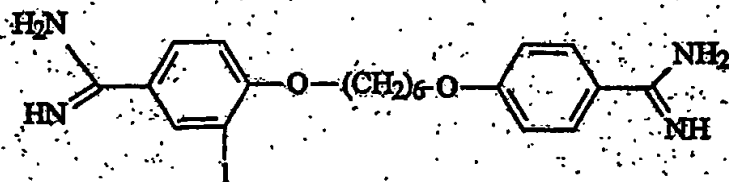
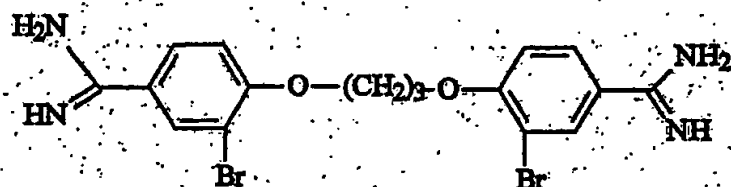
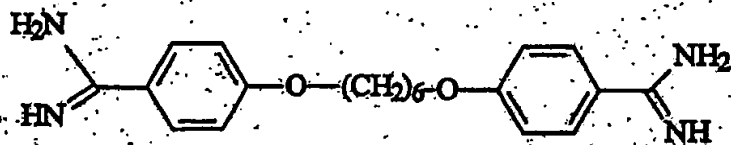
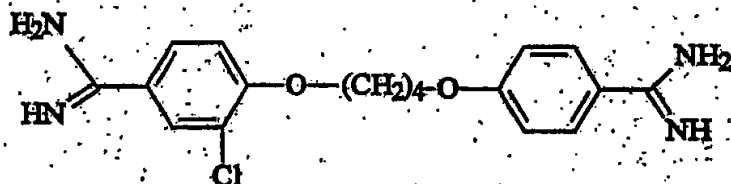


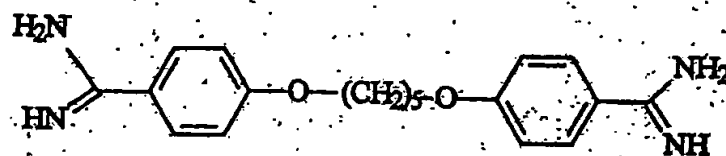
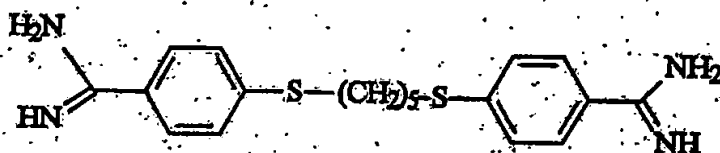
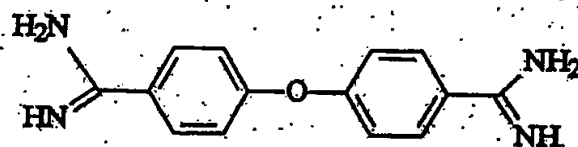
Example III





In another embodiment, the present invention provides a method of inhibiting carcinoma progression wherein matriptase plays a role in a subject in need of such inhibition comprising administering to a subject an effective amount of a compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof. Compounds A, B, C, D, E, F and G have the following respective structures:

COMPOUND ACOMPOUND BCOMPOUND CCOMPOUND D

COMPOUND ECOMPOUND FCOMPOUND G

In another embodiment, the invention provides a method of treating malignancies, pre-malignant conditions, and pathologic conditions in a subject which are characterized by the expression of single-chain (zymogen) and/or two-chain
 5 (activated) form of matriptase comprising administering a therapeutically effective amount of a compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

In yet another embodiment, the invention provides a method of therapy which results in the inhibition of matriptase in a subject in need of such inhibition which
 10 comprises administering a therapeutically effective amount of at least one compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

5 In a further embodiment, the invention provides a method of diagnosing cancer comprising exposing a tissue sample to an antibody or immunogenic fragment thereof which recognizes and binds to a product of matriptase mediated proteolysis of a matriptase substrate.

15 (A) administering to a subject, that is to be tested for a pre-malignant or malignant lesion, or other pathologic condition, which is characterized by the presence of a product of matriptase mediated proteolysis of a substrate of matriptase, a labeled agent which recognizes and binds to the product of matriptase mediated proteolysis; and

20 In still another embodiment, the invention provides an *in vitro* method of diagnosing the presence of a pre-malignant lesion, a malignancy, or other pathologic condition, in a subject, which is characterized by the presence of a product of matriptase mediated proteolysis of a substrate of matriptase comprising the steps of:

(B) exposing the biological sample to a labeled agent which recognizes and binds to the product of matriptase mediated proteolysis; and

- 10 -

4. Brief Description of the Drawings

Figs. 1A and 1B show how matriptase converts single-chain HGF into smaller fragments, which can be recognized by anti- α -chain HGF and β -chain HGF antibodies; Fig. 1A shows the silver-stained protein patterns of HGF incubated overnight without (0) or with increasing amount of matriptase (); Fig. 1B shows Western immunoblot. HGF incubated overnight without (-) or with (+) matriptase was immunoblotted with anti-HGF (α chain) or anti-HGF (β chain);

Figs. 2A and 2B show how matriptase-cleaved HGF stimulates scattering on MDCK cells, plates show wells in the 96-well plates with about 3000 MDCK cells after 20 h in the absence (CRL) or presence of matriptase-cleaved HGF (HGF/MTP) or untreated HGF (HGF) at the dilutions shown, with leupeptin being included in every culture;

Fig. 3 shows how matriptase-cleaved HGF stimulates c-Met tyrosine phosphorylation; A549 cells were treated in the absence of HGF (NO HGF), matriptase-cleaved HGF (HGF/MTP), or HGF alone (HGF) for 5 min at 37 °C; equal amounts of lysed cell protein were immunoprecipitated with anti-c-Met antibody followed by immunodetection using anti-phosphotyrosin antibody (Tyr-p) as described below; in controlling for the amounts of c-Met in each sample, immunoblots were then stripped and detected with anti-c-Met antibody (c-Met); wherein the shown arrow points to the p145 kDa α chain of c-Met under reducing conditions;

Fig. 4 shows how plasminogen is not a substrate for matriptase, this figure shows the silver-stained protein patterns of plasminogen incubated overnight without (0) or with increasing amount of matriptase (); with the highest amount of matriptase used is 8-fold of the lowest amount of matriptase that cleaves HGF;

Figs. 5A and 5B show how Pro-uPA is activated by matriptase cleavage; Fig. 5A shows how single-chain pro-uPA is converted into two-chain form uPA by matriptase; wherein pro-uPA was incubated overnight active matriptase in the absence (0) or presence of increasing amount of matriptase (); the cleaved products were analyzed by electrophoresis followed by silver-staining; and Fig. 5B shows how matriptase cleavage of pro-uPA generates an active protease; pro-uPA was either incubated for 30 min with matriptase (closed circles) or incubated 1 min with plasmin

(opened circles) prior to the assay, plasmin (opened triangles), uPA (closed triangles), and matriptase (closed squares) all exhibit low activity;

Fig. 6 shows a ribbon model 3-D structure of matriptase; the structure is obtained through homology modeling and refinement by molecular dynamics simulation;

Fig. 7 shows a model 3-D structure of matriptase with compound A; the structure is obtained through molecular dynamics docking using the generalized effective potential;

Fig. 8 shows a model 3-D structure of matriptase with compound B; the structure is obtained through molecular dynamics docking using the generalized effective potential

5. Detailed Description of the Invention

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described.

20 Definitions

By "matriptase" is meant a trypsin-like protein, with a molecular weight of between 72-kDa and 92-kDa and is related to SEQ ID NO: 4 or is a fragment thereof. It can include both single-chain and double-chain forms of the protein. The zymogen form (inactive form) of matriptase is a single-chain protein. The two-chain form of matriptase is the active form of matriptase, which possesses catalytic activity. Both forms of matriptase are found to some extent in milk and cancer cells because extracellular matrix (ECM) remodeling is necessary to both normal and pathologic remodeling processes. Both cancer cells and milk contain the different forms of matriptase. However, in milk the dominant form is the activated form of matriptase which then binds to HAI-1.

By "matriptase modulating compound" or "matriptase modulating agent" is meant a reagent which regulates, preferably inhibits the activation of matriptase (*e.g.*,

cleavage of the matriptase single-chain zymogen to the active two-chain moiety) or the activity of the two-chain form of matriptase. This inhibition can be at the transcriptional, translation, and/or post-translational levels. Additionally, modulation of matriptase activity can be via the binding of a compound to the zymogen or
5 activated forms of matriptase.

By "matriptase expressing tissue" is meant any tissue which expresses any form of matriptase, either malignant, pre-malignant, normal tissue, or tissue which is subject to another pathologic condition.

By "malignancy" is meant to refer to a tissue, cell or organ which contains a
10 neoplasm or tumor that is cancerous as opposed to benign. Malignant cells typically involve growth that infiltrates tissue (*e.g.*, metastases). By "benign" is meant an abnormal growth which does not spread by metastasis or infiltration of the tissue. The malignant cell of the instant invention can be of any tissue; preferred tissues are epithelial cells.

By "tumor progression" or "tumor metastasis" is meant the ability of a tumor to
15 develop secondary tumors at a site remote from the primary tumor. Tumor metastasis typically requires local progression, passive transport, lodgement and proliferation at a remote site. This process also requires the development of tumor vascularization, a process termed angiogenesis. Therefore, by "tumor progression" and "metastasis," we
20 also include the process of tumor angiogenesis.

By "pre-malignant conditions" or "pre-malignant lesion" is meant a cell or
tissue which has the potential to turn malignant or metastatic, and preferably epithelial cells with said potential. Pre-malignant lesions include, but are not limited to:
25 atypical ductal hyperplasia of the breast, actinic keratosis (AK), leukoplakia, Barrett's epithelium (columnar metaplasia) of the esophagus, ulcerative colitis, adenomatous colorectal polyps, erythroplasia of Queyrat, Bowen's disease, bowenoid papulosis, vulvar intraepithelial neoplasia (VIN), and dysplastic changes to the cervix.

By "other condition" or "pathologic conditions" is meant any genetic
susceptibility or non-cancerous pathologic condition relating to any disease
30 susceptibility or diagnosis.

By "tumor formation-inhibiting effective amount" is meant an amount of a compound, which is characterized as inhibiting activation of matriptase or matriptase activity, and which when administered to a subject, such as a human subject, prevents the formation of a tumor, or causes a preexisting tumor, or pre-malignant condition, to enter remission. This can be assessed by screening a high-risk patient for a prolonged period of time to determine that the cancer does not arise and/or the pre-malignant condition is reversed. This also can be assessed by imaging of the subject with a tumor to determine whether the mass of the tumor is shrinking. A tumor formation-inhibiting effective amount also includes an amounts which provides ameliorative to relief to the subject. The tumor formation-inhibiting effective amount can also be assessed based on its effect on blood circulation of inhibitors, such as BBIC. Preferred tumor formation-inhibiting effective amounts of agents such as BBIC are in the range of 100 µg/kg to 20 mg/kg body weight of the subject. More preferred ranges include 1 µg/kg to 10 mg/kg body weight of the subject.

By "labeling agent" is meant to include fluorescent labels, enzyme labels, and radioactive labels. By "radiolabel" or "radioactive label" is meant any radioisotope for use in humans for purposes of diagnostic imaging. Preferred radioisotopes for such use include: ^{67}Cu , ^{67}Ga , ^{99}Te , ^{131}I , ^{123}I , ^{125}I , ^{111}In , ^{188}Re , ^{186}Re and ^{90}Y . By "fluorescent label" is meant any compound used for screening samples (e.g., tissue samples and biopsies) which emits fluorescent energy. Preferred fluorescent labels include fluorescein, rhodamine and phycoerythrin.

By "biological sample" is meant a specimen comprising body fluids, cells or tissue from a subject, preferably a human subject. Preferably the biological sample contains cells, which can be obtained via a biopsy or a nipple aspirate, or are epithelial cells. The sample can also be body fluid that has come into contact, either naturally or by artificial methods (e.g. surgical means), a malignant cell or cells of a pre-malignant lesion.

By "matriptase expressing tissue" is meant any biological sample comprising one or more cells which expresses a form or forms of matriptase.

By "subject" is meant an animal, preferably mammalian, and most preferably human.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab fragments and F(ab)₂ fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies (mAb), chimeric antibodies and humanized antibodies. The production of antibodies and the protein structures of complete, intact antibodies, as well as antibody fragments (e.g., Fab fragments and F(ab)₂ fragments) and the organization of the genetic sequences that encode such molecules are well known and are described, for example, in Harlow *et al.*, ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1988).

By "immunogenic fragment" is meant a portion of a matriptase protein which induces humoral and/or cell-mediated immunity but not immunological tolerance.

By "epitope" is meant a region on an antigen molecule (e.g., matriptase) to which an antibody or an immunogenic fragment thereof binds specifically. The epitope can be a three dimensional epitope formed from residues on different regions of a protein antigen molecule, which, in a native state, are closely apposed due to protein folding. "Epitope" as used herein can also mean an epitope created by a peptide or hapten portion of matriptase and not a three dimensional epitope.

The present invention is based on recent discoveries relating to a novel serine protease system, matriptase and its cognate inhibitor (the hepatocyte growth factor activator inhibitor 1). Considering its unique characteristics: the epithelial and cancer cells origin, the membrane integral property, and the multiple putative regulatory domains, the invention is generally based on the hypothesis that matriptase functions as an upstream activator in matrix progression. In particular, the present invention is based on the discovery that matriptase can convert single chain hepatocyte growth factor/scatter factor precursor to the active factor that can induce scatter of Madin-Darby canine kidney epithelial cells and can activate c-Met tyrosine phosphorylation in A549 human lung carcinoma cells. In addition, it has been discovered that matriptase can activate urokinase plasminogen activator but not plasminogen.

While the foregoing detailed description is focused on methods for diagnosing cancer, particularly breast cancer, the subject invention is directed to all conditions wherein matriptase plays a role.

The invention provides novel cancer diagnosis and therapy methods based in part on a novel mechanism in the epithelial-mesenchymal interaction that an epithelial membrane activator regulates the activation of the stromal-origin factors in the extracellular matrix degradation and in the cell migration. These observations provide a possible direct link that has been missing from current model of tissue remodeling.

Identification of a Mechanism Associated With Matriptase Activity

The present inventors have conducted extensive research the initial results of which allowed for the discovery of matriptase and the elucidation of some aspects of its activity. The discoveries associated with matriptase have been incorporated in a research program aimed at providing diagnosis and therapy protocols for the treatment of various cancers, particularly breast cancer. These protocols are based in part on the elucidation of the mechanistic aspects associated with matriptase activity and the utilization of computer modeling tools in combination with biological testing in the identification of antibody diagnostic agents and small molecule matriptase inhibitors as therapeutic agents.

Matriptase is a trypsin-like serine protease with two regulatory modules: two tandem repeats of the complement subcomponent C1r/s domain and four tandem repeats of LDL receptor domain (Lin *et al.*, *J. Biol. Chem.* 274: 18231-6 (1999)). In order to evaluate the role of matriptase in physiological conditions, its expression in human milk was studied. It was found that milk-derived matriptase strongly interacts with fragments of HAI-1 to form SDS-stable complexes.

The mosaic protease is characterized by trypsin-like activity and two regulatory modules (*e.g.*, LDL receptor and complement subcomponent C1r/s domains). Matriptase was initially purified from T-47D human breast cancer cells.

In breast cancer cells, matriptase was detected mainly as an uncomplexed form; however, low levels of matriptase were detected in SDS-stable, 110- and 95-kDa complexes that could be dissociated by boiling. In striking contrast, only the complexed matriptase was detected in human milk.

The complexed matriptase has now been purified by a combination of ionic exchange chromatography and immunoaffinity chromatography. Amino acid

sequences obtained from the matriptase-associated proteins reveal that they are fragments of an integral membrane, Kunitz-type serine protease inhibitor that was previously reported to be an inhibitor (termed HAI-1) of hepatocyte growth factor activator.

- 5 In addition, matriptase and its complexes were also detected in four milk-derived, SV-40 T-antigen-immortalized mammary luminal epithelial cell lines, but not in two human foreskin fibroblasts nor in HT1080 fibrosarcoma cell line. The milk-derived matriptase complexes are likely to be produced by the epithelial components of lactating mammary gland *in vivo*, and the activity and function of matriptase may
10 be differentially regulated by its cognate inhibitor, comparing breast cancer with the lactating mammary gland.

- It has now been discovered that matriptase and its cognate Kunitz-type serine protease inhibitor can be characterized as an extracellular, trypsin-like protease system. *In vivo* expression of matriptase and its inhibitor was observed in breast
15 cancer cells, breast epithelial cells, and in other human carcinomas including ovarian, endometrial, and colon. In no case were matriptase or the inhibitor found in fibroblasts or fibrosarcomas. Matriptase is a type 2 integral membrane serine protease that contains tandem repeats of putative regulatory domains: the LDL receptor domains and the CUB (C1r/s, Uegf and Bone morphogenetic protein-1) domains (2;3)
20 (also see updated sequence in the GenBank/EBI Data Bank with accession number AF118224).

- The cognate inhibitor of matriptase is a type 1 integral membrane protein containing two Kunitz domains separated by a LDL receptor domain. The inhibitor has been characterized as an inhibitor of hepatocyte growth factor activator (HGF).
25 Although originally isolated from breast cancer cell conditioned medium and human milk, both matriptase and the inhibitor are also found on cell surface, mostly by anchoring in the membrane through their transmembrane domains. The soluble form matriptase is likely to be resulted from the shedding of the surface-bound matriptase.

- These characteristics of matriptase suggest that this protease is involved in cell
30 migration and progression by acting as a direct upstream regulator. Without wishing to be bound by any particular theory, it is believed that matriptase is presented on the tumor or epithelial cell surface and interacts with a variety of factors and digests a

broad range of substrates. These actions, together, may induce dramatic changes in the extracellular environment and in the surrounding cells contributing to cell migration, progression, and metastasis.

To further elucidate the biological roles of matriptase, the substrates of this protease were investigated. Particular attention was directed at HGF. HGF, also known as scatter factor (SF), was originally identified as a potent mitogen for mature hepatocytes and as a fibroblast-derived factor with the ability to induce scattering of epithelial colonies. HGF/SF can stimulate cell motility, promote matrix progression, and induce morphogenesis. Although other factors such as blood coagulation factor XIIa, HGF activator, uPA, and tPA have been reported to be able to activate HGF/SF, no epithelial-origin activator has been reported.

Our interest on HGF/SF was further stimulated by the fact that the endogenous kunitz-type inhibitor of matriptase is the HGF activator inhibitor 1 (HAI-1).

In elucidating the mechanisms of matriptase action, the following methods and materials were employed:

Antibodies _ Polyclonal antibodies to hepatocyte growth factor (HGF) _-chain (C-20) and _-chain (N-19) were purchased from Santa Cruz Biotechnology (Calne, Wilshire, U.K). Monoclonal anti-human Met antibodies (clones DL-21 and DL-24) were purchased from Upstate (Lake Placid, NY). The recombinant anti-phosphotyrosine antibody (RC20:HRPO) was from Transduction Laboratory (Lexington, KY).

Cell lines and protein substrates _ Human lung carcinoma cell line A549 was from ATCC. Madin-Darby canine kidney (MDCK II) epithelial cell lines and the single-chain form HGF protein were the generous gifts from Dr. George Vande Woude (NCI, NIH, Fredrick, MD). Single-chain form of human urokinase plasminogen activator (pro-uPA) was purchased from American Diagnostic Inc (Greenwich, CT). Plasminogen, plasmin, and fluorescent substrate peptides N-tert-Butoxy-Carbonyl (N-t-Boc)-Gln-Ala-Arg-7-Amido-4-Methylcoumarin (AMC), N-t-Boc-Leu-Gly-Arg-AMC, N-t-Boc-_ -Benzyl (Bz)-Glu-Gly-Arg-AMC, N-t-Boc-_ -Bz-Glu-Ala-Arg-AMC, N-Succinyl (Suc)-Ala-Phe-Lys-AMC, Suc-Leu-Leu-Val-Tyr-AMC, Suc-Ala-Ala-Pro-Phe-AMC, and Suc-Ala-Ala-Ala-AMC were from Sigma chemical company (St. Louis). All the other chemicals are reagent-graded.

Purification of active Matriptase _ Active matriptase, in the complex form with its endogenous inhibitor HAI-1, was purified from human milk. To separate matriptase from the bound HAI-1, the complex was dissociated by acid and resolved in 10% SDS polyacryamide gel electrophoresis. The proteins were stained by Zinc stain Kit (Bio-Rad, Hercules, CA), the 70-kDa active matriptase was sliced out, and eluted from gel using Electro-Eluter (Bio-Rad, Hercules, CA) under non-denatured condition (Tris-Glycine buffer pH 8.3).

Determination of matriptase activity _ The enzyme activity of matriptase was measured at room temperature in a reaction buffer containing 100 mM Tris-HCl (pH 8.5) and 100 μ g/ml of BSA using fluorescent peptide as substrate. In brief, 10 μ l of enzyme solution and 10 μ l of peptide substrate were added to a cuvette containing 180 μ l of the reaction buffer. The mixture was mixed well, placed back to a fluorescent spectrophotometer (HITACH F4500), and the release of fluorescence resulting from hydrolysis of the peptide substrate was recorded with excitation at 360 nm and emission at 480 nm.

Determination of kinetic parameter _ Substrate concentration versus initial reaction velocity were analyzed by the Michaelis-Menten equation and plotted using SigmaPlot software. Double reciprocal (Lineweaver-Burk) plots thus derived were used to determine Vmax and Km values.

Cleavage of protein substrates _ Single chain HGF protein, plasminogen, or pro-uPA were incubated with various amounts of matriptase in 100 mM Tris-HCl (pH8.5) overnight at room temperature. Incubation was stopped by boiling the mixture in SDS sample buffer. The cleaved products were then separated on SDS-PAGE electrophoresis and analyzed by Western blot hybridization or by silver stain.

Scattering assay _ The MDCK II cell line was maintained at 37 $^{\circ}$ C in Dulbecco's Modified Eagle (DME) medium supplemented with 10% fetal calf serum (FCS). The scatter assay was carried out in the 96-well culture plate (18). To each well was added 150 μ l DME medium supplemented with 5% FCS and leupeptin at 100 μ g per ml; HGF or matriptase-cleaved HGF was added into one well and serial two-fold dilutions were made with sequential 150 μ l aliquots of medium transferred from well to well. About 3,000 MDCK II cells in 150 μ l medium were added to each well and the plate was placed in a 37 $^{\circ}$ C incubator for 20-24 h. Media were removed

and cells were fixed and stained for 15 min with 5% crystal violet in 50% methanol. Cell scattering (spreading and dispersion of epithelial colonies) was examined under light microscopy.

Met phosphorylation detection _ A549 cells were grown confluent in RPMI medium supplemented with 10% FCS. After 3-hour serum-starvation, cells were incubated 5 min incubation at 37 _C with 450 ng/ml HGF or matriptase-cleaved HGF in RPMI medium supplemented with 5% FCS. Where it is needed, leupeptin was included in the medium at 100 _g/ml. Medium were removed, cells were rinsed with 1x PBS and collected by centrifugation following trypsinization. After washed one more time with 1x PBS, cell pellets were frozen in dry-ice. The frozen cell pellets were either stored at -80 _C for later extraction or immediately extracted as described below. Cells were thawed on ice, extracted by suspension in buffer containing 50 mM Tris-HCl (pH7.5), 5 mM EDTA, 150 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM sodium ortovanadate, 100 _g/ml PMSF, 10 _g/ml leupeptine, 10 _g/ml apropeptenin, and 1% Triton X-100. Extracts were clarified by centrifugation for 15 min at 12,000 xg in a microfuge, and the protein concentration was determined by BCATM protein assay kit (PIERCE, Rochford, IL) using bovine serum albumin (BSA) as standard. About 2 mg protein of extracts were immunoprecipitated using anti-c-Met antibody and Pansorbin (Calbiochem, La Jolla, CA). The protein-antibody-Pansorbin immunocomplex was collected by centrifugation, washed twice with extraction buffer then with 1x PBS, and then dissociated by boiling in SDS-sample buffer. The Pansorbin was removed by centrifugation, the supernatant fractions were subjected to 8% SDS electrophoresis, and the proteins were detected by Western immunoblot with anti-phosphotyrosin antibody. The same immunoblot were then stripped in 100 mM 2-mercaptoethonal, 2% SDS, and 62.5 mM Tris-HCl (pH7.6) for 30 min at 50 _C and re-probed with anti-c-Met antibody.

The mechanistic investigation of matriptase action provided the following results:

Comparison of the amino acid sequences of the protease domain of matriptase with that of other serine proteases revealed that matriptase contains a trypsin-like substrate specificity pocket, and a Asp residue positioned at the bottom of the

specificity pocket. To investigate in detail the substrate specificity of matriptase, we compared the K_m and V_{max} of matriptase to a variety of protease substrate peptides.

Table 1. Kinetic parameters of matriptase for various peptide substrates.

10 10^{-4} of matriptase solution and 10^{-4} of peptide substrate were added to a cuvette containing 180μ l reaction buffer, the release of fluorescence resulting from hydrolysis of the peptide was recorded in a fluorescent spectrophotometer at room temperature with excitation at 360 nm and emission at 480 nm. K_m and V_{max} were determined by double reciprocal plot derived from Michaelis-Menten equation:

PeptideSubstrate	P4-P3-P2-P1-Amc	K_m (μ M)	V_{max} (nM Amc/min)
2	Glu-Ala-Arg ^a	3.808	76.34
3	Leu-Gly-Arg ^a	13.6	308.9
4	Gln-Gly-Arg	33.52	527.5
5	Glu-Gly-Arg	47.5	169.8
6	Ala-Phe-Lys	69.9	523.6
7	Leu-Leu-Val-Tyr ^a	^b	^b
8	Ala-Ala-Pro-Phe ^a	^b	^b
9	Ala-Ala-Ala ^a	^b	^b

10

^a Gln-Ala-Arg is a standard substrate for trypsin, Leu-Gly-Arg is a substrate for uPA, Leu-Leu-Val-Tyr and Ala-Ala-Pro-Phe are the substrates for chymotrypsin, and Ala-Ala-Ala is the substrate for elase.

^b No cleavage activity was detected with these substrates at concentration of 200μ M.

15

Table 1 shows that, as expected from the sequence analysis, matriptase, like trypsin, prefers to cleave peptides at Arg or Lys. Not surprisingly, the best substrate peptide for matriptase is Boc-Gln-Ala-Arg-Amc, the best substrate for trypsin. No released fluorescence was detected from the substrate for chymotrypsin or elastase (table 1, peptide substrates 7, 8, 9). Matriptase appears to prefer to bind to peptides containing small side chain amino acids, such as Ala and Gly, at P2 site (table 1, peptide substrates 1-5). Peptides containing P2 Ala are better substrates for matriptase than peptides containing P2 Gly (compare peptides 1, 2 with peptides 3-5). The binding affinity of matriptase to the formers is about 30-fold higher than to the latter. Interestingly, a change from Gln to Glu at the P3 site significantly reduces the V_{max} (compare peptide 1 with 2) without causing significant change on K_m . Information from matriptase peptide substrate specificity led us to look for

25

macromolecular substrates that require cleavage at Arg for activation. HGF/SF is secreted as an inactive, single chain precursor by stromal cells, and it is activated by proteolytic conversion to the two-chain form factor by cleavage at Arg-495 (19) in the extracellular environment. Furthermore, HAI1 is also the inhibitor of matriptase.

5 HGF/SF appears to be an ideal substrate candidate for matriptase.

To determine if HGF/SF can be cleaved by matriptase, inactive HGF/SF purified from the condition medium of fibroblast cells cultured in the absence of serum (a generous gift from Dr. Vande Woude in NCI) was taken in the following experiments. Fig. 1A shows that this HGF/SF preparation is primarily composed of the single-chain form protein about 97 kDa on SDS-PAGE (lane a). There are some
10 minor species of 64 kDa and 33 kDa apparent molecular weight corresponding to the expected sizes of activation cleaved α - and β - chains of HGF/SF, respectively. After incubation with the active form matriptase, the amount of 64 kDa and 33 kDa molecules increased at the expense of 97 kDa molecule. The lowest amount of
15 matriptase for HGF cleavage observed was at a molar ratio proximally 500:1 for HGF to matriptase. Western immunoblot analysis using monoclonal antibodies to α - and β - chain HGF shows that the 64 kDa species contains α -HGF and the 33 kDa species contains β -HGF (Fig. 1B). The anti- α -HGF antibody apparently does not recognize the single-chain form HGF/SF (Fig. 1B). Thus, matriptase can convert the single chain
20 HGF/SF into fragments containing α - and β - HGF.

To further investigate if the cleavage of HGF/SF by matriptase corresponds to a process of activation, we performed a well-characterized assay specific to HGF activity, the MDCK cell scatter assays. The single-chain HGF/SF was first incubated with matriptase to allow complete cleavage, the cleaved products were then diluted in
25 the medium and incubated with MDCK II cells for 20-24 h at 37 °C. To eliminate the activation of HGF/SF by other activators in the serum, leupeptin was included in every scattering assay. It was noticed that leupeptin can also inhibit matriptase enzyme activity (unpublished observation from our laboratory).

Figs. 2A and 2B show the images of plates of MDCK II cells after incubation
30 with various dilutions of untreated (HGF) or matriptase-treated (HGF/MTP) HGF/SF. In the control sample (CRL), cells aggregated into tight clusters after 20-24 h culture. In the presence of 1:4096 dilution of HGF/MTP, cell islands were more spread and

there was a high proportion of single cells. In contrast, at the same dilution of HGF, cells were not significantly different from the control cells. The scatter morphology was more and more profound with progressively lower dilutions of HGF/MTP. At 1:512 dilution of HGF/MTP, cells completely separated from each other and many of them exhibited elongated fibroblast-like shapes. On the other hand, scattering morphology induced by the untreated HGF appeared at much lower dilution (Figs. 2A and 2B, HGF, 1:512). The residual scattering activity in these experiments apparently come from the residual active HGF in the HGF/SF preparation (Fig. 1). Cells grown in the presence of matriptase alone retained the morphology as those in the control (data not shown). Cells incubated with untreated HGF/SF in the absence of leupeptin were examined as a positive control for HGF activation by serum; the scattering activity of HGF/SF in these samples was as expected (data not shown). These observations showed that the scattering induced by HGF/MTP is indeed contributed by the matriptase activation of HGF.

HGF operates its function through binding to its cell surface receptor, Met. Upon binding to the HGF, Met is activated following phosphorylation at tyrosine residues. In Fig. 3, c-Met activation was examined in the A549 human lung carcinoma cell line. Like the experiments for Figs. 2A and 2B, leupeptin was also included in these treatments. Tyrosine phosphorylated c-Met was observed in the cells incubated with matriptase-cleaved HGF (HGF/MTP) and with untreated HGF (HGF), but not in the cells incubated in the absence of HGF (No HGF). Phosphotyrosine of c-Met in cells incubated with untreated HGF appears to be caused by the residual active HGF contamination in the preparation. Leupeptin did not affect the total c-Met expression, the c-Met phosphorylation, or the total pattern of tyrosine phosphorylation (data not shown).

Plasminogen shares high homology with HGF and its activation also requires a cleavage at Arg. It seemed likely that plasminogen would be a substrate of matriptase as well. However, to our surprise, matriptase failed to cleave plasminogen. As shown in Fig. 4, plasminogen remained as a 94 kDa single-chain form even in the presence of matriptase at 8-fold higher than that required to cleave HGF.

Plasmin has long been regarded as the enzyme that converts pro-uPA to active uPA. However, the level of active uPA is not reduced in the urine of mice bearing a

targeted disruption of the plasminogen gene, suggesting the existence of plasmin independent pro-uPA activation. Plasma kallikrein, trypsin-like proteases from human ovarian tumors, a T cell-associated serine protease, cathepsins B and L, nerve growth factor α , human mast cell trypase, and prostate specific antigen have also been reported to activate pro-uPA. However, the roles of these enzymes *in vivo*, relevant to activation of pro-uPA *in vivo* are not clear. We were therefore interested to see if matriptase can activate pro-uPA. Fig. 5A showed that after incubation with the matriptase, the 55 kDa single-chain pro-uPA was converted into smaller fragments. One of these cleavage products clearly appeared on the protein gel as the 33 kDa molecule, which resembles the size of the active uPA protease (Fig. 5A). The cleaved product exhibited enzymatic activity toward the fluorescent peptide substrate, Boc-Leu-Gly-Arg-AMC, for uPA (Fig. 5B, compare the closed circles with closed triangles). This activity was not derived from matriptase, since matriptase alone only exhibited negligible background activity (Fig. 5B, closed squares). The same amounts of pro-uPA produced a similar activity after cleavage by plasmin (Fig. 5B, opened circles).

By using the active matriptase isolated from human milk, we report in this study that matriptase cleaves and converts HGF into a biologically functional factor that can induce c-Met activation of and stimulate epithelial cells scattering. In addition, we also noted that matriptase can activate pro-uPA but not plasminogen. These results further support matriptase to be an upstream regulator in matrix progression. Most significantly, these results reveal a novel mechanism in the control of tissue remodeling that involve an upstream epithelial membrane activator and downstream stromal effectors.

Tissue remodeling is a process observed both in physiological processes such as embryonic development, morphogenesis, and wound healing, and in pathological situations such as cancer progression and metastasis. Two essential changes occur during these processes: 1) epithelial-mesenchymal transition transforms rigid epithelial cells to the mobile mesenchymal cell that can migrate distances; 2) extracellular matrix degradation that open path for the migrating cells. HGF is a potent inducer of epithelial-mesenchymal transition; engagement of HGF to its epithelial compartment c-Met triggers various signaling subsequently induces various

cellular responses. HGF is secreted as inactive precursor by stromal cells and is proteolytically activated in the extracellular environment. Therefore, activation of HGF/SF needs to occur in the close vicinity of the epithelial cells. The network of the serine proteases uPA/uPA-receptor/plasminogen and the zinc-dependent metalloproteinases have been proposed to be responsible for the majority of proteolysis of pericellular proteins. However, both systems are largely synthesized by the stromal cells and require indirect mechanisms for their recruitment and activation on the surfaces of epithelial cells. Thus, an epithelial-derived protease like matrilysin would preferably provide the missing link in this process.

10 In addition, the process of extracellular matrix degradation requires protease cascades; proteolytic activation of one protease depends on the proteolytic activation of its activator. A central question in this process is that of whether there is a primary initiating enzyme. The finding that matrilysin can activate both pro-uPA and HGF/SF suggesting that it should be considered as a central regulator in epithelial cell remodeling and/or progression of extracellular matrix. Matrilysin contains a serine protease activation motif; its activation is likely to require the proteolytic cleavage at Arg-Val. Autoactivation has been reported with the recombinant protease domain of matrilysin. These observations suggest that activation of matrilysin may occur intermolecularly. As we have shown previously, matrilysin contains multiple non-catalytic domains that involve in protein-protein interaction. These domains may assist in accumulating other factors at the contact sites between cancer and stromal cells; these domains may also be involved in the activation of matrilysin. Clues may lie in its N-terminal non-catalytic domains, particularly the CUB domains. It has been shown that CUB-EGF module in the C1r and C1s subunits play essential role in the assembly and auto-activation of the C1 complex. The same domains in matrilysin may also be critical for its function. Specifically, they may contribute to an intermolecular interaction between matrilysin molecules. Thus these domains may be involved in the auto-activation of matrilysin, like they do in C1 complex.

15 Matrilysin appears to have selectivity for the macromolecular substrates. In our experiments, matrilysin did not cleave plasminogen, despite that the high sequence homology between plasminogen and HGF. Another group recently has also reported this substrate selectivity. In the same report, it was also showed that matrilysin/MT-

SP1 has selectivity for a basic residue at P3 or P4 site. The sequence at the activation cleavage site of HGF and plasminogen is P4-(Lys)-P3-(Gln)-P2-(Gly)-P1-(Arg) and P4-(Cys)-P3-(Pro)-P2-(Leu)-P1-(Arg), respectively. Thus lacking of P3 or P4 basic residue in plasminogen might attribute to its not being cleaved by matriptase.

- 5 Alternatively, there might be conformational differences between HGF and plasminogen. The kringle domains of plasminogen contain a Lys-binding site that serves to mediate its localization to fibrin and to cellular surfaces. Plasminogen circulates in the blood in a globular and closed conformation; upon binding to the surface, it shifts to an extended and opened conformation. This conformation change
10 promotes its recognition by its activator and its rapid convert to plasmin. The kringle domains on HGF/SF also contribute to its binding to its receptor c-Met. However, cleavage-activation of HGF/SF does not depend on its binding to c-Met. It is possible that the single chain form HGF/SF exhibit a more open conformation than does plasminogen; and that matriptase can distinguish this subtle structural difference.

15

Identification of Small Molecules Capable of Inhibiting Matriptase Action

- As disclosed *supra*, the focus of the subject invention is to provide therapies based on compounds capable of interfering with matriptase protease activity. In one aspect, the invention provides therapies based on existing compounds which are
20 identified through computational modeling as inhibitors of matriptase activity. The compounds are identified through structure-based three-dimensional (3D) database searching. The compounds identified through database searching are processed through biological tests to identify one or more lead compounds for clinical testing and/or rational drug design refinement.

- 25 Computationally predicting a compound's binding affinity to a host protein involves utilizing the three dimensional (3-D) structures of the host protein and the compound. The 3-D structure of the compound is obtained from a database of chemical compounds. The 3-D structure of the host protein can also be obtained from a protein database.

- 30 The invention provides potent and specific matriptase inhibitors through a structure-based drug discovery approach. The methodology employed in the discovery of matriptase inhibitors is disclosed in U.S. Patent Application No.

09/301,339, filed on April 29, 1999, the contents of which are hereby incorporated by reference in their entirety.

Briefly, in structure-based 3D-database searching for drug discovery, once the 3D structure of the target molecule (a receptor or an enzyme) is determined, large
5 chemical databases containing the 3D structures of hundreds of thousands of structurally diverse synthetic compounds and natural products are searched through computerized molecular docking to identify small molecules that can interact effectively with the target or host molecule.

In spite of the massive increase in the number of biological molecules whose 3-
10 D structure has been elucidated, the majority of proteins of known primary structure (amino acid sequence) do not have a known tertiary (or 3-D) structure.

For drug design involving target proteins of unknown tertiary structure, a model structure can be constructed based on the known tertiary structure of a protein which is homologous to the target protein. The structure of the homologous protein is used
15 to construct a template structure of all or part of the target protein. The structure obtained through homology modeling provides a working structure for further refinement. The working structure for the protein not having a known structure is obtained by refining the template structure.

In forming a template 3-D structure of the host protein, each atom of the
20 backbone of the target protein is assigned a position corresponding to the position of the equivalent backbone atom in the homologous protein. Similarly, each atom of a side chain of the target protein having an equivalent side chain in the homologous protein is assigned the position corresponding to the position of the atom in the equivalent side chain of the homologous protein. The atom positions of a side chain
25 not having an equivalent in the homologous protein are determined by constructing the side chain according to preferred internal coordinates and attaching the side chain to the backbone of the host protein.

The template structure thus obtained is refined by minimizing the internal energy of the template protein. During the refinement, the positions of the atoms of
30 the side chains having no equivalents in the homologous protein are adjusted while keeping the rest of the atoms of the template protein in a fixed position. This allows the atoms of the constructed side chains to adapt their positions to the part of the

template structure determined by homology. The full template structure is then minimized (relaxed) by allowing all the atoms to move. Relaxing the template 3-D structure of the protein eliminates unfavorable contacts between the atoms of the protein and reduces the strain in the template 3-D structure.

5 Based on the refined structure of the target protein, a host-guest complex is formed by disposing a compound from a compound database in a receptor site of the protein. The structure of the host-guest complex is defined by the position occupied by each atom in the complex in a three dimensional referential.

10 A geometry-fit group is formed by selecting the compounds which can be disposed in the target binding site without significant unfavorable overlap with the atoms of the protein. For each compound in the geometry fit group, a predicted binding affinity to the receptor site of the host protein is determined by minimizing an energy function describing the interactions between the atoms of the compound and those of the protein. The minimization of the energy function is conducted by
15 changing the position of the compound such that a guest-host complex structure corresponding to a minimum of the energy function is obtained. The compounds having the most favorable energy interaction with the atoms of the binding site are identified for optional further processing, for example through display and visual inspection of compound protein complexes to identify the most promising compound
20 candidates.

25 The displayed complexes are visually examined to form a group of candidate compounds for in vitro testing. For example, the complexes are inspected for visual determination of the quality of docking of the compound into the receptor site of the protein. Visual inspection provides an effective basis for identifying compounds for
in vitro testing.

After putative binding compounds have been identified, the ability of such compounds to specifically bind to matriptase is confirmed *in vitro* and/or *in vivo*.

Potent and selective inhibitors are tested further in their ability to inhibit colony-formation in soft-agarose. Compounds having good *in vitro* activity are tested
30 *in vivo*. Tumor bearing mice are treated with therapy, based on the compounds and the effect on the tumor size is observed. Compounds showing effective tumor reduction are then used in clinical trial protocols.

Computational identification of compounds having potential matriptase inhibitory activity

5 To date, the experimental 3D structure matriptase has not been determined. However, the structures of a number of other similar proteases have been determined through X-ray crystallography.

10 The X-ray structure of human thrombin, entry 1hxe from the protein databank was used as template for building the 3D structure of matriptase through homology modeling. Since the sequence identity and similarity between template and model are 34% and 53%, respectively, the 3D structure of matriptase can be modeled accurately. By analogy to other serine proteases, the serine protease domain (b-chain) of matriptase has a catalytic triad formed by Ser191, His42 and Asp97, that is positioned on the surface of the protease. A negatively charged residue, Asp185, is located at the bottom of the S1 binding site, Fig. 6, which is consistent with the experimental data showing the preference of matriptase for substrates with positively charged residues, Arg/Lys at the P1 position.

15 Since the S1 site is considered to be the primary binding site in serine proteases, it is a good target for inhibitor design. The disulfide bond between Cys216 and Cys187 stabilize the position of Asp185. Gly215, Cys216, Ala217 and Gln218 are at the entrance of the S1 binding pocket. An anionic site formed by Asp46, Asp47 and Asp91 may also be important for inhibitor binding to matriptase. A putative hydrophobic S1' binding site is marked by Leu18, Ala20, Leu21, Ile26 and Trp58. The disulfide bond between Cys27 and Cys43 may stabilize the position of Ile26, which may be important for the geometry of this binding site.

25 The sequence for MTP was obtained from sequencing data. Templates for homology modeling were obtained by searching the Protein Databank, using the program BLAST. As discussed above, the structure of thrombin, entry 1HXB with 34% identities, 53% similarity and 6% gaps, was used as a template for modeling MTP using the program MODELLER. The structure obtained from homology modeling was refined using the molecular dynamics program CHARMM. Hydrogen atoms were assigned to the modeled structure using the program HBUILD from

CHARMM. Then it was solvated by inserting it in a 30 Å sphere of water and deleting solvent molecules with heavy atoms closer than 2.5 Å to any protein heavy atom. This was followed by minimization to relieve bad contacts and 100 ps molecular dynamics simulation. The refined structure of MTP was used for structure-based database

5 search.

Structure-Based Database Search. The refined structure of MTP was used as the target for structure-based 3D database search. The program DOCK was used for computer aided database screening to identify potential inhibitors. Shape and binding energy scoring functions were used to screen and rank ligands. Filters were used to
10 eliminate molecules that have more than 10 flexible bonds to avoid considering overly flexible molecules, also molecules with fewer than 10 or more than 50 heavy atoms were discarded.

The screening of the large NCI database was done in two stages with the ligand flexibility being considered in both. In the first one, two minimization cycles with 50
15 iterations maximum were considered for each compounds from the database. The best scoring 10,000 molecules were considered in the second stage when 100 minimization cycles and 100 maximum iterations per cycle were done to refine the position of the ligand and its score. The top 2,000 compounds were considered for selecting potential inhibitors for matriptase by inspecting if they contain ionizable groups that will bind
20 to the S1 site and the anionic site. After the third screening 69 compounds were selected for biological testing.

The refined structure of matriptase obtained from molecular dynamics simulation (see Fig. 6) was used for screening the NCI database. The S1 site, the anionic site and the hydrophobic S1' site were used to prepare the template for
25 docking with DOCK. In order to save time, the docking was done in two rounds. In the first one a crude minimization was used and the best scoring 10,000 compounds were considered for the second round with longer minimization. The best scoring 2,000 compounds were considered and 69 were selected for testing.

Molecular dynamics simulation with the generalized effective potential was
30 used for flexible ligand docking in explicit solvent. This method was chosen because it was shown to have better conformational sampling than simple molecular dynamics. The inhibitor was positioned perpendicular to the line between the S1 site

and the anionic site. The starting distance between the two amidine groups and Asp185 from the S1 site was 15 – 22 Å. The active site of MTP-inhibitor complex was solvated by centering the molecule around Ser191 and the center of the inhibitor. Then a 20 Å radius water sphere was added for solvation. Water molecules closer than 2.5 Å to any protein or inhibitor heavy atom were deleted. The solvated structure was minimized and equilibrated at 300 K. The docked structure was obtained after several ns molecular dynamics simulation with the generalized effective potential.

Biological Activity

To confirm the efficacy of the subject analogs in inhibiting matriptase activity. Active matriptase was purified from human milk as will be described later. Active urokinase-type plasminogen activator (uPA) was purified by aminobenzamidine-Sepharose 6B (Amersham Pharmacia, Piscataway, NJ) from a partially purified uPA from human urine (a gift from a local company in Taiwan). Bovine α -trypsin, bovine thrombin, and fluorescent peptide substrates N-tert-Butoxy-Carbonyl (N-t-Boc)-Gln-Ala-Arg-7-Amido-4-Methylcoumarin (AMC), N-t-Boc-Leu-Gly-Arg-AMC, and N-t-Boc-Leu-Arg-Arg-AMC were purchased from Sigma (Sigma Chemical Company, St. Louis).

Activated matriptase, in the complex form with its endogenous inhibitor HAI-1, was purified from human milk as described previously. To separate the activated matriptase from the bound HAI-1, the complex was dissociated by acid, and resolved in 10% SDS polyacrylamide gel electrophoresis. The proteins were stained by Zinc stain Kit (Bio-Rad, Hercules, CA). Gels containing the 70-kDa active matriptase was sliced out, and eluted from gel using Electro-Eluter (Bio-Rad, Hercules, CA) under non-denatured condition (Tris-Glycine buffer pH 8.3). Purified active matriptase was then stored at -80 °C in acid solution.

Inhibitory activity of the compound inhibitors against each protease was measured at room temperature using fluorescent substrate peptides in 100 mM Tris-HCl (pH 8.5) containing 100 µg/ml of bovine serum albumin. To a cuvette containing 170 µl buffer was added 10 µl of enzyme solution and 10 µl of inhibitors. After pre-incubation, 10 µl of substrate was added and the solution was mixed well by shaking the cuvette. The residue enzyme activity was then determined by following the change of fluorescence released by hydrolysis of the fluorescent substrates in a fluorescent

spectrophotometer (HITACH F4500) with excitation at 360 nm and emission at 480 nm. Peptide N-t-Boc-Gln-Ala-Arg-AMC was used as substrate for matriptase and trypsin, peptide N-t-Boc-Leu-Gly-Arg-AMC was used as substrate for uPA, and peptide N-t-Boc-Leu-Arg-Arg-AMC was used as substrate for thrombin.

5 The inhibitory activity of each compound was first investigated under the same concentration to a fixed amount of matriptase. Compound that exhibited inhibition was then further analysis for its K_i value using Dixon plotting. We recorded the rate of hydrolysis in the presence of 6-7 different concentrations of each inhibitor. A straight line of the concentration of inhibitor versus the reciprocal values of the rate of
10 hydrolysis was plotted with SigmaPlot software. Two lines were obtained from two unsaturated substrate concentrations; the X value of the intersection of these lines gives the value of $-K_i$.

Based on the results obtained with the compounds identified to date, it is anticipated that these compounds will comprise significant therapeutic application. In
15 particular, these compounds should inhibit matriptase proteolytic activity and the biological effects associated therewith. More specifically, these compounds should selectively inhibit matriptase carcinoma progression, cell proliferation and/or differentiation and/or promote apoptosis especially of cancer and other neoplastic cells.

20

Table 2: Initial screening of compound inhibitors for matriptase.

Over 95% inhibition	15
Increase activity	3

Table 2 shows the results from an initial screening. Among 69 compounds examined, 44 could inhibit at least 60% of the matriptase activity; 15 of them actually
25 inhibited over 95% of the activity. There were 6 compounds not useful in this study: 3 compounds had high absorbency which made them incompatible with our activity

assay; other three compounds, instead of reducing, increased the rate of releasing fluorescence by matriptase hydrolysis. From the 15 compounds that inhibited more than 95% of matriptase activity 5 were analyzed further for their K_i value as described in Methods.

5

TABLE 3

Compound	Matriptase	Trypsin	uPa	Thrombin
A	191 nM	672 nM	1.98 μ M	796 nM
B	535 nM	393 nM	1.57 μ M	946 nM
C	942 nM	3,80 μ M	14.4 μ M	224 nM
D	3.93 μ M	10.5 μ M	-	-
E	1.155 μ M	-	-	-
F	-	-	-	-
G	-	-	-	-

Table 3 shows the K_i of five most potent compounds. Compound A, that is the most potent is also the most selective when tested against matriptase, trypsin, uPA and thrombin. Compound C, that is the unsubstituted analog of compound A is about 5 times less potent against matriptase but is about 3 times more potent inhibitor of thrombin. As a conclusion an iodine at position 2 can improve the potency and selectivity of the inhibitor for matriptase. It has to be noted that the length of A, 19 Å, is comparable to the distance between Asp185 and the anionic site, that is about 20 Å. This data also supports our model. A comparison of C and E suggests that shorter linker between the two benzamidine groups decreases the potency of the inhibitor. Halogens in position 2 of the aromatic ring increase the potency of the inhibitor. The only compound identified through database search that has two halogens is compound B. This compound is about 2.5 times less potent than A and has little selectivity between matriptase, trypsin and thrombin.

Docking results. Compounds A and B were docked into the active site of MTP using molecular dynamics simulation with the generalized effective potential,

See Figs. 7 and 8, respectively. This method takes into account besides ligand flexibility, the receptor flexibility and solvent effects. We have found that both molecules can bridge between the S1 site and the anionic site. Both ligands interact with Asp185 from the S1 site and Asp91 from the anionic site. Despite the fact that
5 compound B is shorter it can still bridge between the two sites because the flexibility of the protein allows a conformational change that results in a shorter distance between the anionic site and Asp185.

Thus, the compounds produced according to the invention will be used to treat conditions wherein inhibition of matriptase is therapeutically beneficial. This will
10 include conditions that involve abnormal cell growth and/or differentiation such as cancers and other neoplastic conditions. Examples of cancers which may be treated according to the invention include colon, pancreatic, prostate, head and neck, gastric, renal, brain and CML.

The subject therapies will comprise administration of at least one compound
15 according to the invention in an amount sufficient to elicit a therapeutic response, e.g., inhibition of carcinoma progression.

The compound may be administered by any pharmaceutically acceptable means, by either systemic or local administration. Suitable modes of administration include oral, dermal, e.g., via transdermal patch, inhalation, via infusion, intranasal, rectal,
20 vaginal, topical parenteral (e.g., via intraperitoneal, intravenous, intramuscular, subcutaneous, injection).

Typically, oral administration or administration via injection is preferred. The subject compounds may be administered in a single dosage or chronically dependent upon the particular disease, condition of patient, toxicity of compound, and whether
25 this compound is being utilized alone or in combination with other therapies. Chronic or repeated administration will likely be preferred based on other chemotherapies.

The subject compounds will be administered in a pharmaceutically acceptable formulation or composition. Examples of such formulations include injectable solutions, tablets, milk, or suspensions, creams, oil-in-water and water-in-oil
30 emulsions, microcapsules and microvesicles.

These compositions will comprise conventional pharmaceutical excipients and carriers typically used in drug formulations, e.g., water, saline solutions, such as

phosphate buffered saline, buffers, surfactants.

The subject compounds may be free or entrapped in microcapsules, in colloidal drug delivery systems such as liposomes, microemulsions, and macroemulsions.

Suitable materials and methods for preparing pharmaceutical formulations are

- 5 disclosed in *Remington's Pharmaceutical Chemistry*, 16th Edition, (1980). Also, solid formulations containing the subject compounds, such as tablets, and capsule formulations, may be prepared.

- Suitable examples thereof include semipermeable materials of solid hydrophobic polymers containing the subject compound which may be in the form of
10 shaped articles, e.g., films or microcapsules, as well as various other polymers and copolymers known in the art.

- The dosage effective amount of compounds according to the invention will vary depending upon factors including the particular compound, toxicity, and inhibitory activity, the condition treated, and whether the compound is administered alone or
15 with other therapies. Typically a dosage effective amount will range from about 0.0001 mg/kg to 1500 mg/kg, more preferably 1 to 1000 mg/kg, more preferably from about 1 to 150 mg/kg of body weight, and most preferably about 50 to 100 mg/kg of body weight.

- The subjects treated will typically comprise mammals and most preferably will
20 be human subjects, e.g., human cancer subjects.

- The compounds of the invention may be used alone or in combination. Additionally, the treated compounds may be utilized with other types of treatments, e.g., cancer treatments. For example, the subject compounds may be used with other chemotherapies, e.g., tamoxifen, taxol, methothrexate, biologicals, such as antibodies,
25 growth factors, lymphokines; or radiation, etc. Combination therapies may result in synergistic results.

The preferred indication is cancer, especially the cancers identified previously.

- While the invention has been described in terms of preferred embodiments, the skilled artisan will appreciate that various modifications, substitutions, omissions and
30 changes may be made without departing from the spirit thereof. Accordingly, it is intended that the scope of the present invention be limited solely by the scope of the following claims, including equivalents thereof.

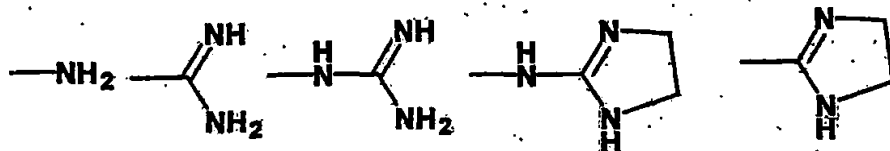
WHAT IS CLAIMED IS:

1. A method of inhibiting carcinoma progression wherein matriptase plays a role in a subject in need of such inhibition comprising administering to a subject an effective amount of a compound comprising two groups capable of being positively charged at physiological pH, which can be the same or different, said groups being
- 5 linked by a chemical moiety having a length of between 5 and 30 Å.

2. The methods of Claim 1, wherein the chemical moiety has a length of between 15 and 24 Å.

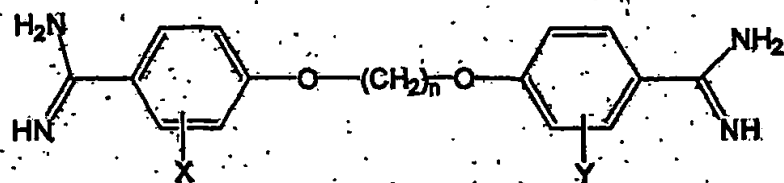
10

3. The method of Claim 1, wherein the positively charged groups are individually selected from the following groups:

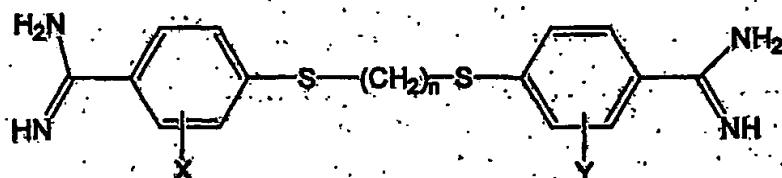


- 15 4. The method of Claim 1 wherein the compound is selected from the compounds of Examples I through VI which have the structures below, wherein X and Y are any substituents that are not detrimental to inhibition of matriptase

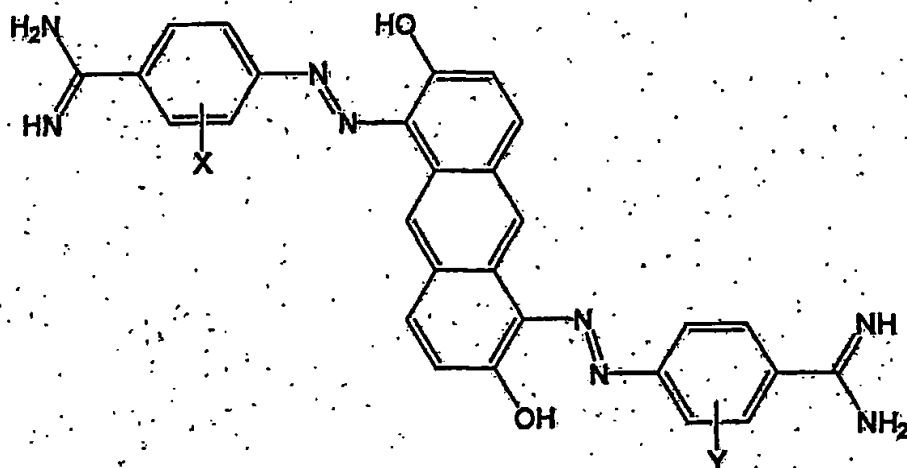
Example I

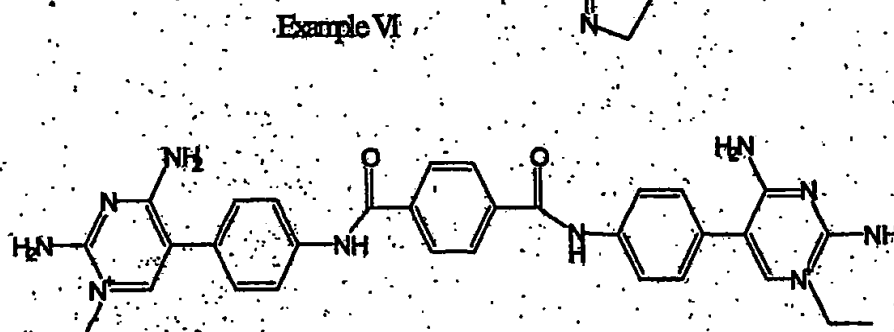
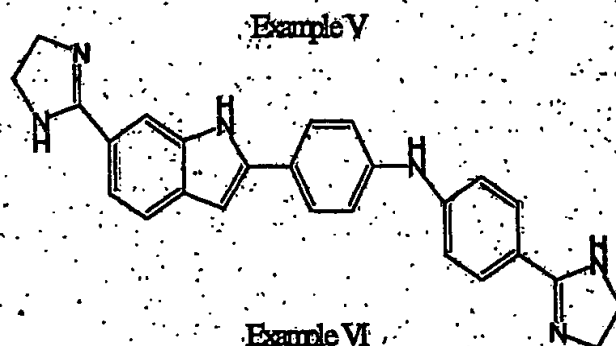
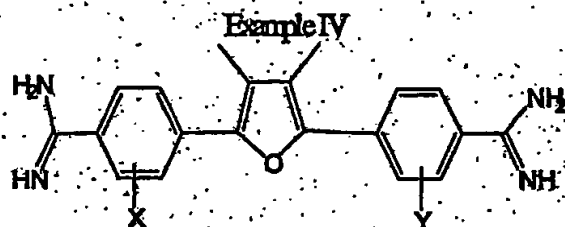


Example II



Example III





5. A method of inhibiting carcinoma progression wherein matriptase plays a role in a subject in need of such inhibition comprising administering to a subject an effective amount of a compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

6. The method of Claim 5, wherein compound A to G is administered to a subject in a daily dose of between 0.1 and 500 mg for each kilogram of the subject's weight.

7. The method of Claim 6, wherein compound A is administered to a subject in a daily dose of about 1.0-200 mg for each kilogram of the subject's weight.

8. The method of Claim 5, wherein said compound inhibits carcinoma progression related protease cascade.

9. The method of Claim 8, wherein inhibiting protease cascade comprises
5 inhibiting matriptase activity.

10. The method of Claim 9, wherein inhibiting matriptase activity comprises inhibiting activation of a substrate involved in carcinoma progression.

10 11. The method of Claim 10, wherein the substrate is an extracellular matrix component, a growth factor or a protease.

12. The method of Claim 10, wherein the substrate is HGF/SF or pro-uPA.

15 13. The method of Claim 8, wherein inhibiting protease cascade comprises inhibiting the production of a protein fragment selected from the group consisting of the catalytically active fragment of HGF/SF produced by cleavage of HGF/SF and uPA produced by cleavage of pro-uPA.

20 14. Method of therapy which results in the inhibition of matriptase in a subject in need of such inhibition which comprises administering a therapeutically effective amount of at least one compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

25 15. The method of Claim 13, wherein said compound antagonizes carcinoma progression related protease cascade.

16. The method of Claim 13, wherein said method results in the inhibition
of carcinoma progression.

30

17. A method of treating cancer comprising administering a therapeutically effective amount of at least one compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

5 18. The method of Claim 16, wherein said cancer is selected from the group consisting of breast, ovarian, prostate, endometrial, colon, pancreatic, head and neck, gastric, renal, brain and CML.

10 19. The method of Claim 16, wherein said compound is administered by a method selected from the group consisting of oral, intranasal, intraperitoneal, intravenous, intramuscular, intratumoral, rectal, and transdermal.

20. The method of Claim 16, wherein the administered amount of said compound ranges from 0.01 mg to 200 mg/kg of the weight of said subject.

15

21. The method of Claim 16, wherein the amount of said compound more preferably ranges from 0.5 mg to 50 mg/kg of the weight of said subject.

20 22. The method of Claim 16, which further comprises the administration of another anticancer compound, radiation, or a compound that induces apoptosis.

23. A pharmaceutical composition which comprises a therapeutically effective amount of a compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof, and a pharmaceutically acceptable carrier.

25

24. The composition of Claim 22, which is suitable for administration via injection, orally, transdermally, intranasally, intraocularly, or rectally.

25. A method of treating malignancies, pre-malignant conditions, and
30 pathologic conditions in a subject which are characterized by the expression of single-

chain (zymogen) and/or two-chain (activated) form of matriptase comprising administering a therapeutically effective amount of a compound selected from the group consisting of compounds A, B, C, D, E, F, G and analogs thereof.

5 26. The method of Claim 24, wherein the malignancy and pre-malignant condition is a condition of the breast.

 27. The method of Claim 22, wherein the pre-malignant condition is selected from the group consisting of: atypical ductal hyperplasia of the breast, actinic
10 keratosis (AK), leukoplakia, Barrett's epiethlium (columnar metaplasia) of the esophagus, ulcerative colitis, adenomatous colorectal polyps, erythroplasia of Queyrat, Bowen's disease, bowenoid papulosis, vulvar intraepithelial neoplasia (VIN), and displastic changes to the cervix.

15 28. A method of diagnosing cancer comprising exposing a tissue sample to an antibody or immunogenic fragment thereof which recognizes and binds to a product of matriptase mediated proteolysis of a matriptase substrate.

 29. The method of Claim 27, wherein the matriptase substrate is uPA or
20 HGF/SF.

 30. The method of Claim 27, wherein the product is active urokinase or active HGF fragment obtained by cleavage of HGF/SF.

25 31. An *in vivo* method of diagnosing the presence of a pre-malignant lesion, a malignancy or other pathologic condition in a subject comprising the steps of:
 (A) administering to a subject, that is to be tested for a pre-malignant or malignant lesion, or other pathologic condition, which is characterized by the presence of a product of matriptase
30 mediated proteolysis of a substrate of matriptase, a labeled agent which recognizes and binds to the product of matriptase mediated proteolysis; and

(B) imaging the subject for the localization of the labeled agent.

32. The method of Claim 30, wherein the labeled agent is an antibody.

5 33. The method of Claim 31, wherein the labeled antibody is a labeled monoclonal antibody.

34. The method of Claim 32, wherein the agent is labeled with a radiolabel or a fluorescent label.

10

35. The method of Claim 33, wherein the radiolabel is selected from the group consisting of: ^{62}Cu , ^{99}Te , ^{131}I , ^{123}I , ^{111}In , ^{90}Y , ^{188}Re , and ^{186}Re .

15 36. An *in vitro* method of diagnosing the presence of a pre-malignant lesion, a malignancy, or other pathologic condition, in a subject, which is characterized by the presence of a product of matriptase mediated proteolysis of a substrate of matriptase comprising the steps of:

- 20 (A) obtaining a biological sample from a subject that is to be tested for a pre-malignant lesion, a malignancy, or other pathologic condition;
- (B) exposing the biological sample to a labeled agent which recognizes and binds to the product of matriptase mediated proteolysis; and
- 25 (C) determining whether said labeled agent bound to the biological sample.

37. The method of Claim 35, wherein the biological sample is a sample comprising epithelial cells.

30 38. The method of Claim 35, wherein the labeled agent is a labeled antibody.

39. The method of Claim 37, wherein the labeled antibody is labeled with a radioisotope or a fluorescent label.

40. The method of Claim 38, wherein the radioisotope is selected from the
5 group consisting of: ^{62}Cu , ^{99}Te , ^{131}I , ^{123}I , ^{111}In , ^{90}Y , ^{188}Re , and ^{186}Re .

1/9

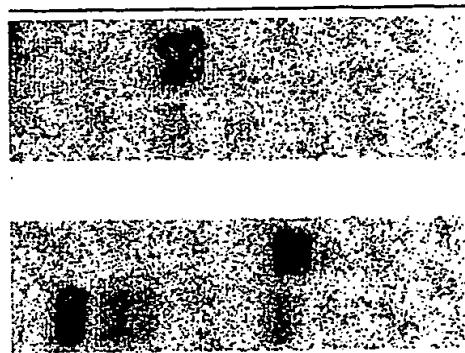


FIG. 1B

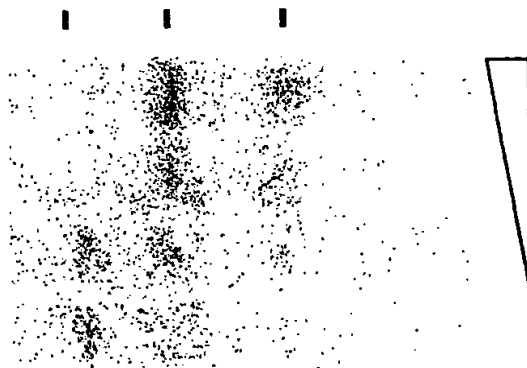


FIG. 1A

2/9

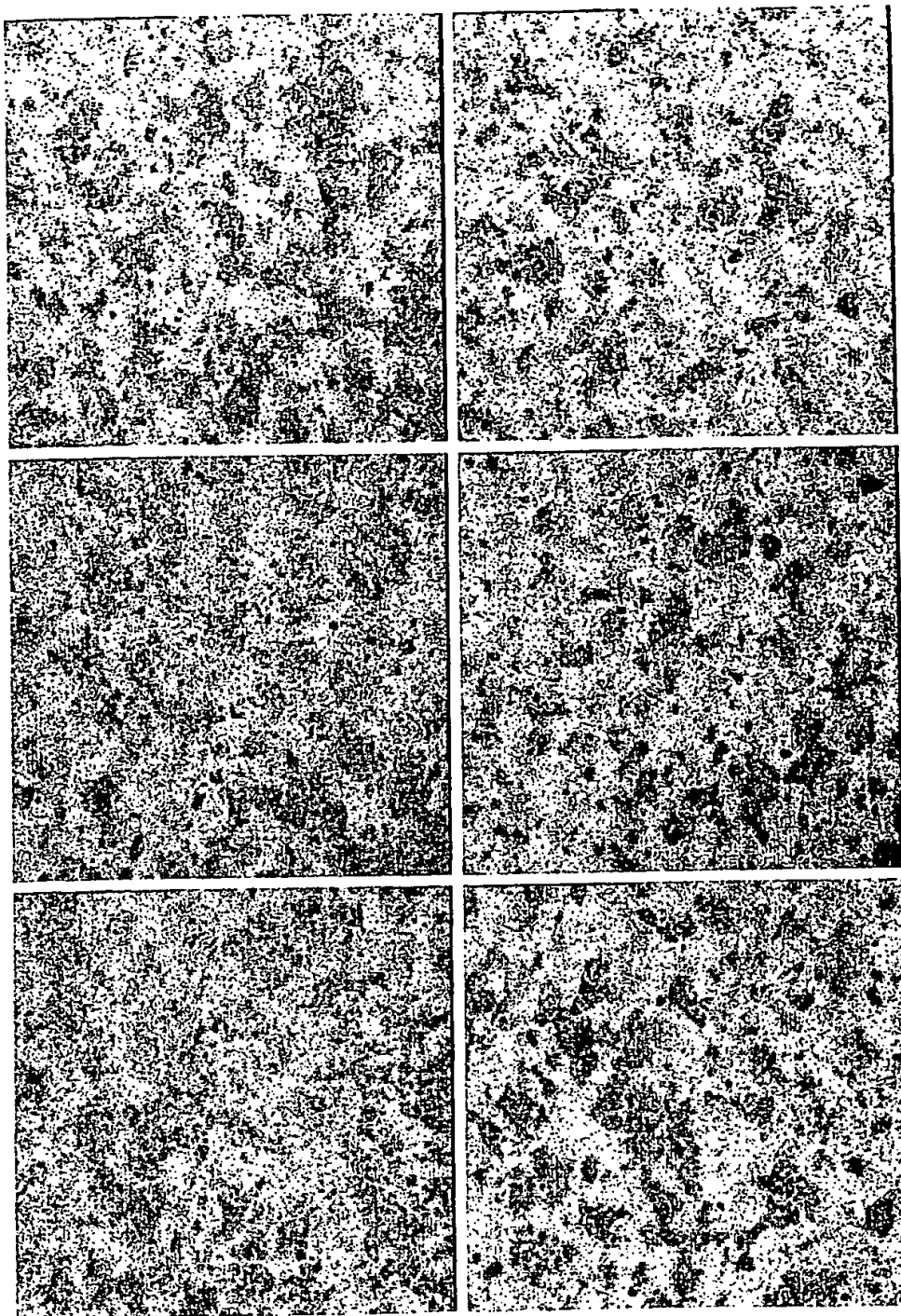


FIG. 2A

SUBSTITUTE SHEET (RULE 26)

3/9

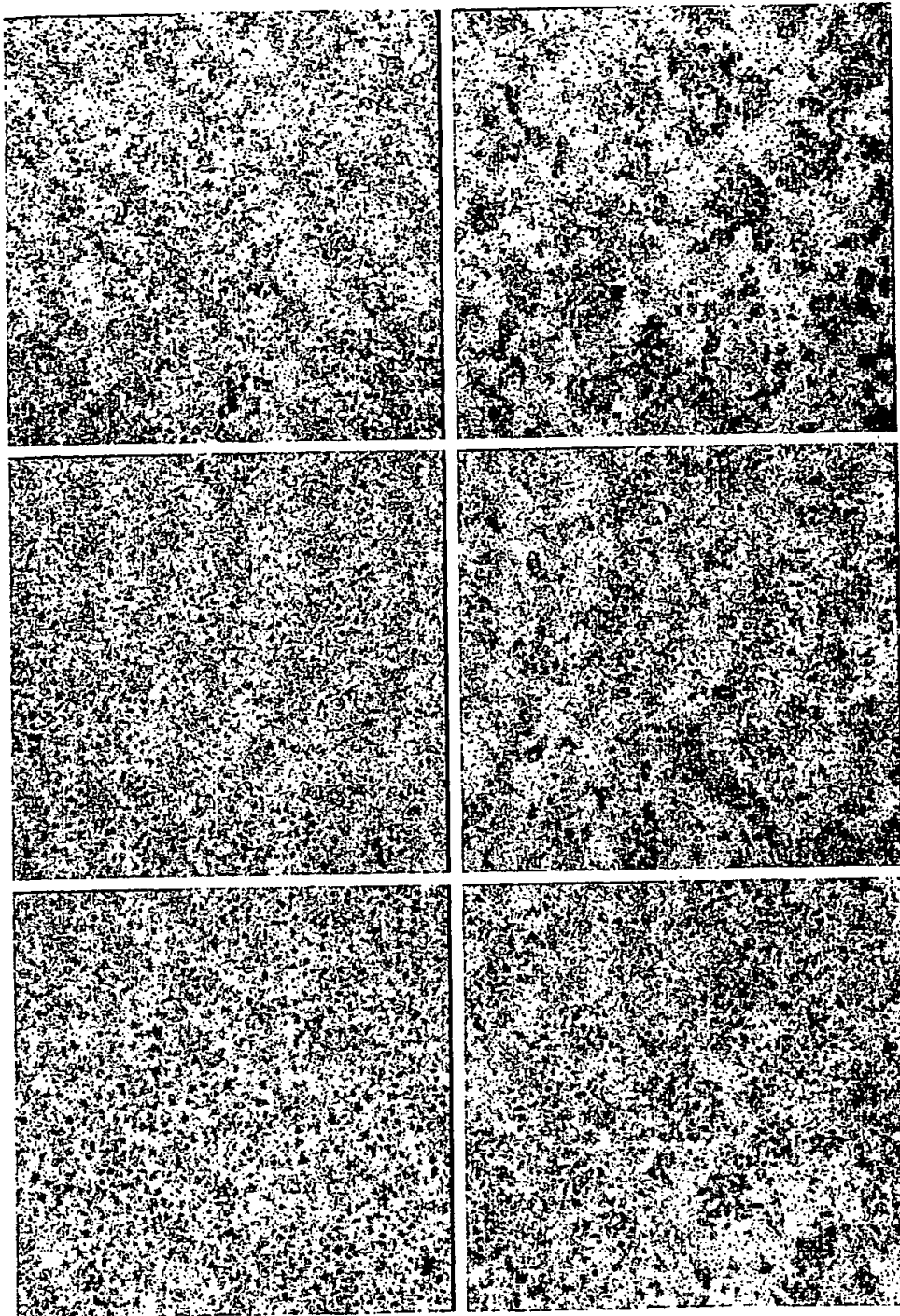


FIG. 2B

SUBSTITUTE SHEET (RULE 26)

4/9

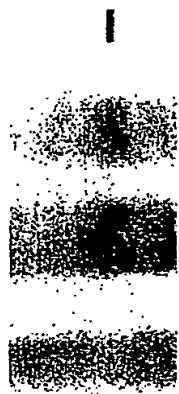


FIG. 3A

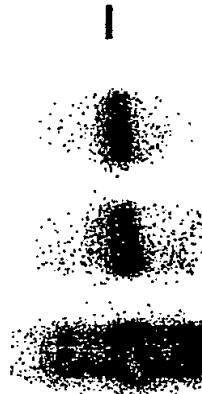


FIG. 3B

5/9

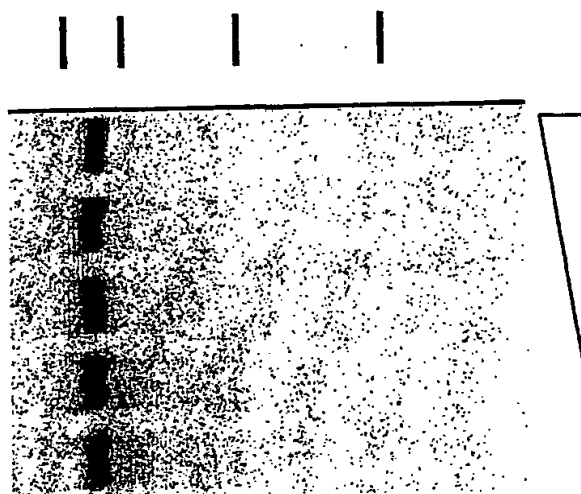
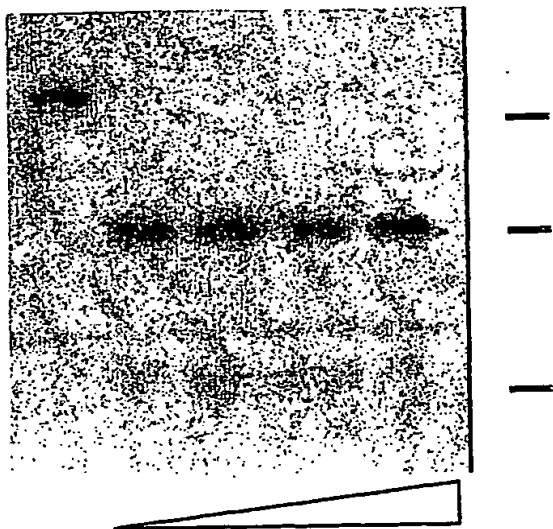
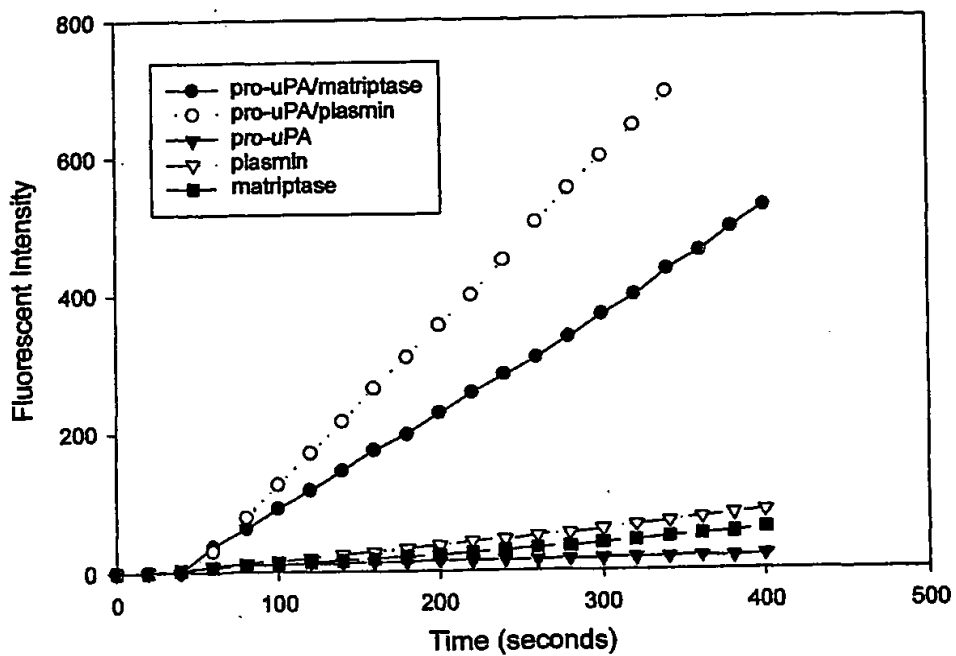
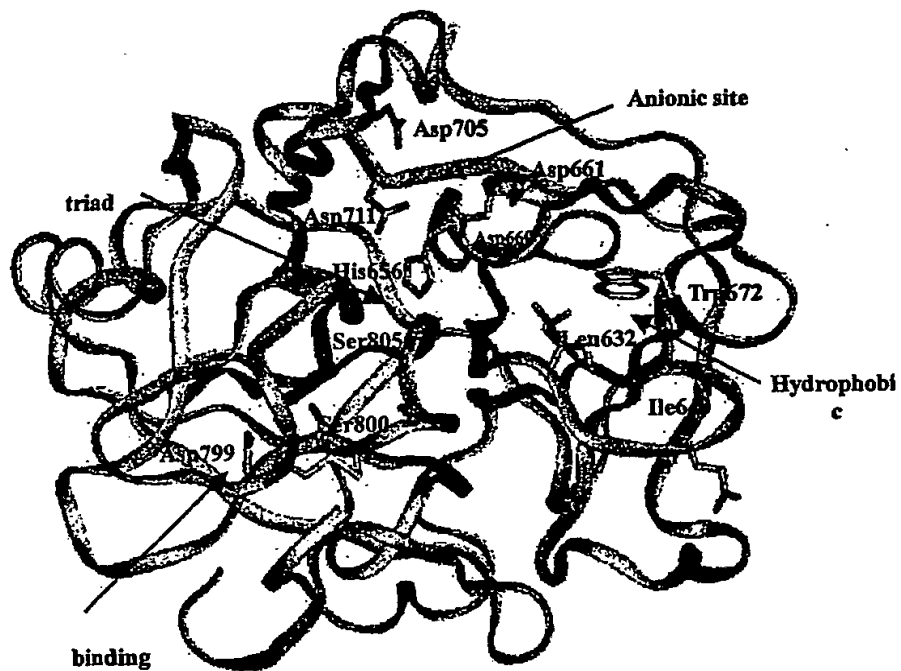


FIG. 4

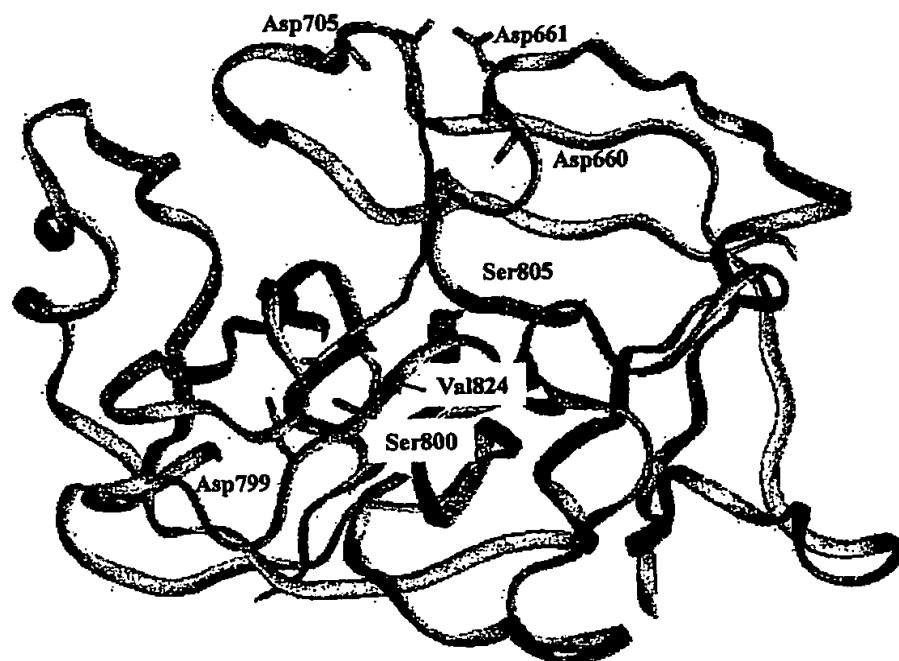
6/9

**FIG. 5A****FIG. 5B**

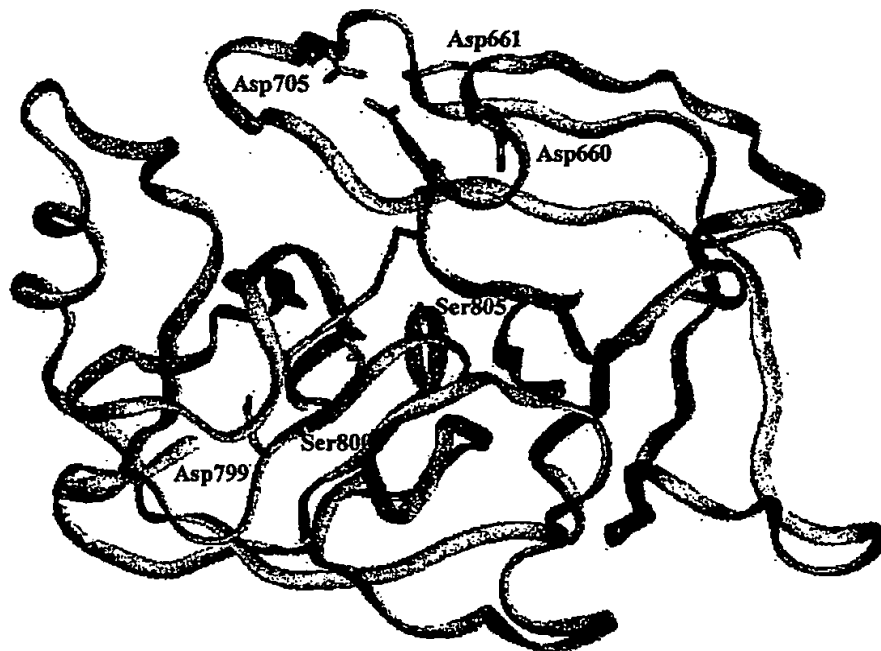
7/9

**FIG. 6**

8/9

**FIG. 7**

9/9

**FIG. 8**